SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Greene, Marianne B.
                     Blumberg, Bruce

    (ii) TITLE OF INVENTION: Human Peroxisome Proliferator
                              Activated Receptor Gamma:
                              Compositions and Methods

    (iii) NUMBER OF SEQUENCES: 8

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Arnold, White & Durkee
        (B) STREET: 321 North Clark Street, Suite 800
        (C) CITY: Chicago
        (D) STATE: IL
        (E) COUNTRY: USA
        (F) ZIP: 60610

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER: US UNKNOWN
        (B) FILING DATE: 08-OCT-1993
        (C) CLASSIFICATION:

    (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: Coughlin, Daniel F.
        (B) REGISTRATION NUMBER: 36,111
        (C) REFERENCE/DOCKET NUMBER: ARCD098

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: 312-744-0090
        (B) TELEFAX: 312-755-4489

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1844 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 179..1603

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CCGACCTTAC CCCAGGCGGC CTTGACGTTG GTCTTGTCGG CAGGAGACAG CACCATGGTG      60

GGTTCTCTCT GAGTCTGGGA ATTCCCGAGC CCGAGCCGCA GCCGCCGCCT GGGGGGCTTG     120

GGTCGGCCTC GAGGACACCG GAGAGGGGCG CCACGCCGCC GTGGCCGCAG AAATGACC       178
```

. . . .

```
ATG GTT GAC ACA GAG ATC GCA TTC TGG CCC ACC AAC TTT GGG ATC AGC          226
Met Val Asp Thr Glu Ile Ala Phe Trp Pro Thr Asn Phe Gly Ile Ser
 1               5                   10                  15

TCC GTG GAT CTC TCC GTA ATG GAA GAC CAC TCC CAC TCC TTT GAT ATC          274
Ser Val Asp Leu Ser Val Met Glu Asp His Ser His Ser Phe Asp Ile
                20                  25                  30

AAG CCC TTC ACT ACT GTT GAC TTC TCC AGC ATT TCT ACT CCA CAT TAC          322
Lys Pro Phe Thr Thr Val Asp Phe Ser Ser Ile Ser Thr Pro His Tyr
            35                  40                  45

GAA GAC ATT CCA TTC ACA AGA ACA GAT CCA GTG GTT GCA GAT TAC AAG          370
Glu Asp Ile Pro Phe Thr Arg Thr Asp Pro Val Val Ala Asp Tyr Lys
        50                  55                  60

TAT GAC CTG AAA CTT CAA GAG TAC CAA AGT GCA ATC AAA GTG GAG CCT          418
Tyr Asp Leu Lys Leu Gln Glu Tyr Gln Ser Ala Ile Lys Val Glu Pro
 65                  70                  75                  80

GCA TCT CCA CCT TAT TAT TCT GAG AAG ACT CAG CTC TAC AAT AAG CCT          466
Ala Ser Pro Pro Tyr Tyr Ser Glu Lys Thr Gln Leu Tyr Asn Lys Pro
                85                  90                  95

CAT GAA GAG CCT TCC AAC TCC CTC ATG GCA ATT GAA TGT CGT GTC TGT          514
His Glu Glu Pro Ser Asn Ser Leu Met Ala Ile Glu Cys Arg Val Cys
            100                 105                 110

GGA GAT AAA GCT TCT GGA TTT CAC TAT GGA GTT CAT GCT TGT GAA GGA          562
Gly Asp Lys Ala Ser Gly Phe His Tyr Gly Val His Ala Cys Glu Gly
            115                 120                 125

TGC AAG GGT TTC TTC CGG AGA ACA ATC AGA TTG AAG CTT ATC TAT GAC          610
Cys Lys Gly Phe Phe Arg Arg Thr Ile Arg Leu Lys Leu Ile Tyr Asp
            130                 135                 140

AGA TGT GAT CTT AAC TGT CGG ATC CAC AAA AAA AGT AGA AAT AAA TGT          658
Arg Cys Asp Leu Asn Cys Arg Ile His Lys Lys Ser Arg Asn Lys Cys
145                 150                 155                 160

CAG TAC TGT CGG TTT CAG AAA TGC CTT GCA GTG GGG ATG TCT CAT AAT          706
Gln Tyr Cys Arg Phe Gln Lys Cys Leu Ala Val Gly Met Ser His Asn
                165                 170                 175

GCC ATC AGG TTT GGG CGG ATC GCA CAG GCC GAG AAG GAG AAG CTG TTG          754
Ala Ile Arg Phe Gly Arg Ile Ala Gln Ala Glu Lys Glu Lys Leu Leu
            180                 185                 190

GCG GAG ATC TCC AGT GAT ATC GAC CAG CTG AAT CCA GAG TCC GCT GAC          802
Ala Glu Ile Ser Ser Asp Ile Asp Gln Leu Asn Pro Glu Ser Ala Asp
            195                 200                 205

CTC CGT CAG GCC CTG GCA AAA CAT TTG TAT GAC TCA TAC ATA AAG TCC          850
Leu Arg Gln Ala Leu Ala Lys His Leu Tyr Asp Ser Tyr Ile Lys Ser
            210                 215                 220

TTC CCG CTG ACC AAA GCA AAG GCG AGG GCG ATC TTG ACA GGA AAG ACA          898
Phe Pro Leu Thr Lys Ala Lys Ala Arg Ala Ile Leu Thr Gly Lys Thr
225                 230                 235                 240

ACA GAC AAA TCA CCA TTC GTT ATC TAT GAC ATG AAT TCC TTA ATG ATG          946
Thr Asp Lys Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu Met Met
                245                 250                 255

GGA GAA GAT AAA ATC AAG TTC AAA CAC ATC ACC CCC CTG CAG GAG CAG          994
Gly Glu Asp Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln Glu Gln
            260                 265                 270

AGC AAA GAG GTG GCC ATC CGC ATC TTT CAG GGC TGC CAG TTT CGC TCC          1042
Ser Lys Glu Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe Arg Ser
            275                 280                 285
```

```
GTG GAG GCT GTG CAG GAG ATC ACA GAG TAT GCC AAA AGC ATT CCT GGT    1090
Val Glu Ala Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile Pro Gly
        290                 295                 300

TTT GTA AAT CTT GAC TTG AAC GAC CAA GTA ACT CTC CTC AAA TAT GGA    1138
Phe Val Asn Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys Tyr Gly
305                 310                 315                 320

GTC CAC GAG ATC ATT TAC ACA ATG CTG GCC TCC TTG ATG AAT AAA GAT    1186
Val His Glu Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn Lys Asp
                325                 330                 335

GGG GTT CTC ATA TCC GAG GGC CAA GGC TTC ATG ACA AGG GAG TTT CTA    1234
Gly Val Leu Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu Phe Leu
                340                 345                 350

AAG AGC CTG CGA AAG CCT TTT GGT GAC TTT ATG GAG CCC AAG TTT GAG    1282
Lys Ser Leu Arg Lys Pro Phe Gly Asp Phe Met Glu Pro Lys Phe Glu
            355                 360                 365

TTT GCT GTG AAG TTC AAT GCA CTG GAA TTA GAT GAC AGC GAC TTG GCA    1330
Phe Ala Val Lys Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp Leu Ala
        370                 375                 380

ATA TTT ATT GCT GTC ATT ATT CTC AGT GGA GAC CGC CCA GGT TTG CTG    1378
Ile Phe Ile Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly Leu Leu
385                 390                 395                 400

AAT GTG AAG CCC ATT GAA GAC ATT CAA GAC AAC CTG CTA CAA GCC CTG    1426
Asn Val Lys Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln Ala Leu
                405                 410                 415

GAG CTC CAG CTG AAG CTG AAC CAC CCT GAG TCC TCA CAG CTG TTT GCC    1474
Glu Leu Gln Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu Phe Ala
            420                 425                 430

AAG CTG CTC CAG AAA ATG ACA GAC CTC AGA CAG ATT GTC ACG GAA CAC    1522
Lys Leu Leu Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr Glu His
            435                 440                 445

GTG CAG CTA CTG CAG GTG ATC AAG AAG ACG GAG ACA GAC ATG AGT CTT    1570
Val Gln Leu Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met Ser Leu
        450                 455                 460

CAC CCG CTC CTG CAG GAG ATC TAC AAG GAC TTG TAC TAGCAGAGAG         1616
His Pro Leu Leu Gln Glu Ile Tyr Lys Asp Leu Tyr
465                 470                 475

TCCTGAGCCA CTGCCAACAT TTCCCTTCTT CCATTTGCAC TATTCTGAGG GAAAATCTGA  1676

CCATAAGAAA TTTACTGTGA AAAAGCGTTT TAAAAAGAAA AGGGTTTAGA ATATGATCTA  1736

TTTTATGCAT ATTGTTTATA AAGACACATT TACAATTTAC TTTTAATATT AAAAATTACC  1796

ATATTATGAA ATTGCAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAA              1844
```

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 476 amino acids
      (B) TYPE: amino acid
      (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Val Asp Thr Glu Ile Ala Phe Trp Pro Thr Asn Phe Gly Ile Ser
1               5                   10                      15
```

```
          Ser Val Asp Leu Ser Val Met Glu Asp His Ser His Ser Phe Asp Ile
                       20                  25              30

          Lys Pro Phe Thr Thr Val Asp Phe Ser Ser Ile Ser Thr Pro His Tyr
                       35                  40              45

    5     Glu Asp Ile Pro Phe Thr Arg Thr Asp Pro Val Val Ala Asp Tyr Lys
                   50                  55              60

          Tyr Asp Leu Lys Leu Gln Glu Tyr Gln Ser Ala Ile Lys Val Glu Pro
          65                      70                  75                  80

   10     Ala Ser Pro Pro Tyr Tyr Ser Glu Lys Thr Gln Leu Tyr Asn Lys Pro
                           85                  90                      95

          His Glu Glu Pro Ser Asn Ser Leu Met Ala Ile Glu Cys Arg Val Cys
                       100                 105             110

   15     Gly Asp Lys Ala Ser Gly Phe His Tyr Gly Val His Ala Cys Glu Gly
                   115                 120             125

          Cys Lys Gly Phe Phe Arg Arg Thr Ile Arg Leu Lys Leu Ile Tyr Asp
                   130                 135             140

          Arg Cys Asp Leu Asn Cys Arg Ile His Lys Lys Ser Arg Asn Lys Cys
          145                 150                 155                 160

   20     Gln Tyr Cys Arg Phe Gln Lys Cys Leu Ala Val Gly Met Ser His Asn
                           165                 170                 175

          Ala Ile Arg Phe Gly Arg Ile Ala Gln Ala Glu Lys Glu Lys Leu Leu
                       180                 185             190

          Ala Glu Ile Ser Ser Asp Ile Asp Gln Leu Asn Pro Glu Ser Ala Asp
                       195                 200             205

   25     Leu Arg Gln Ala Leu Ala Lys His Leu Tyr Asp Ser Tyr Ile Lys Ser
                   210                 215             220

          Phe Pro Leu Thr Lys Ala Lys Ala Arg Ala Ile Leu Thr Gly Lys Thr
          225                 230                 235                 240

   30     Thr Asp Lys Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu Met Met
                       245                 250             255

          Gly Glu Asp Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln Glu Gln
                       260                 265             270

          Ser Lys Glu Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe Arg Ser
                   275                 280             285

   35     Val Glu Ala Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile Pro Gly
                   290                 295             300

          Phe Val Asn Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys Tyr Gly
          305                 310                 315                 320

          Val His Glu Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn Lys Asp
   40                     325                 330             335

          Gly Val Leu Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu Phe Leu
                       340                 345             350

          Lys Ser Leu Arg Lys Pro Phe Gly Asp Phe Met Glu Pro Lys Phe Glu
                   355                 360             365
```

```
      Phe Ala Val Lys Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp Leu Ala
          370             375             380

      Ile Phe Ile Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly Leu Leu
      385             390             395                     400

  5   Asn Val Lys Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln Ala Leu
              405             410                     415

      Glu Leu Gln Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu Phe Ala
              420             425                     430

 10   Lys Leu Leu Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr Glu His
              435             440                     445

      Val Gln Leu Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met Ser Leu
          450             455             460

      His Pro Leu Leu Gln Glu Ile Tyr Lys Asp Leu Tyr
      465             470             475
```

(2) INFORMATION FOR SEQ ID NO:3:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 41 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

    ACTAGTGCGG CCGCCTAGGC CTCGATTTT TTTTTTTTTT T      41

(2) INFORMATION FOR SEQ ID NO:4:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 16 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (ix) FEATURE:
             (A)    NAME/KEY:   Y
             (B)    LOCATION:   3, 12 and 21
             (C)    IDENTIFICATION METHOD:   Y =   C or T

        (ix) FEATURE:
             (A)    NAME/KEY:   R
             (B)    LOCATION:   6, and 15
             (C)    IDENTIFICATION METHOD:   R  =   A or G

        (ix) FEATURE:
             (A)    NAME/KEY:   N
             (B)    LOCATION:   9 and 18
             (C)    IDENTIFICATION METHOD:   N =   G, A, C or T

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TGYGARGGNT GYAARGGNTT YTT
23


(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Cys Glu Gly Cys Lys Gly Phe Phe
1           5


(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ACGTGACCTT TGTCCTGGT    19

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 6 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TGACCT     6


(2) INFORMATION FOR SEQ ID NO:8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 15 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Thr Ile Arg Leu Lys
1          5          10        15